

WP-17

Metadata management and mining Data Continuum

Laura R, Tibor S, Nicola B, Franck S, Jean-François P, ...

- Survey of the different metadata catalogue solutions (Dspace, Invenio, CKAN, Zentity ...)
 - Most of them dedicated either to Library, OpenData mvt (regions, city, ...) or link to GRID technology
 - Nice features (**Reporting/Data Mining**, large deployment, basket, ...)
- Installation of ICAT ongoing in the Lab
 - Try to catch up with latest release (DESY,ESRF,ILL)
 - ICAT 3.4 running at ILL (in house dev, web interface)

- Milestone report mostly complete (due by the end of Sep)
 - Clarification of specifications
 - Catalogue Survey
 - Architecture
- Complete the installation in the lab
 - Ingestion
 - Web portal
- Develop missing ICAT/TopCAT functionalities
 - **Integrate Report / Data Mining interface into ICAT**
 - Mechanism for making users data public
 - Enable users to administrate the authorization
 - Logging made available to the users

- Study of the different Persistent Id solutions (Handle – DOI, PURL, ARK) and providers (DataCite , CrossRef, ...)
 - Review and Test of the APIs
 - Review of contract
- DOI and DataCite chosen by CERN and ILL
- ILL is signing the contract
- Request from DataCite APIs for harvesting publication where Lab experimental data are cited in the form of DOI.
(DataCite summer meeting)

	Handle	DOI	PURL{Z}	ARK
Global scope	++	++	++	++
Persistence	++	++	++	++
Scalability	++	++	+	++
Independence	++	++	++	++
Encoding	+	++	-	+
Transcribability	-	++	-	+
Metadata	++	++	++	++
Opacity	~	~	~	~+
Centralization	+	++	-	++
Security	++	++	++	++
Providers	+	++	+	++
Market penetration	+	++	-	-
Libraries support	++	++	-	+

- Milestone not fully ready.
- Implementation of DOIs for data id
- Harvest publications citing DOI Facility data